Structural Studies of β-Galactosidase, DEAD-box Proteins, and Lysozymes from Phages T4 and P22

The Effect of Cryocooling on Crystals of *E. coli* β-Galactosidase

D.H. Juers and B.W. Matthews
Institute of Molecular Biology, Department of Physics,
and Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403

E. coli β-galactosidase is a relatively large enzyme with an extensive history in biochemistry and molecular biology. The structure was first determined to 2.5 Å resolution using a monoclinic crystal form[1]. Subsequent analysis in an orthorhombic form in concert with cryogenic techniques and the high intensity x-ray source at ALS BL 5.0.2 has improved the resolution to 1.7 Å^2. Further analysis of inhibitor complexes and mutants up to 1.4 Å resolution has produced a basic outline for the reaction coordinate of the enzyme.

An unexpected result from the high resolution studies involves the effect of cryocooling on the crystal packing. The structures determined at ALS (which were done at cryogenic temperature) consistently showed large cell dimension decreases (up to 10 Å) in comparison with room temperature structures. Further investigation showed that the changes are reversible and appear to be due to lattice repacking

We have found two effects which may be responsible for the repacking. First, on cooling, the area of the protein surface involved in lattice contacts increases by 50% (Figure 1). There are substantial alterations in intermolecular contacts, these changes being dominated by the long, polar side chains. For entropic reasons such side chains tend to be somewhat disordered at room temperature but can form extensive hydrogen-bonded networks on cooling. Second, macroscopic density measurements on the vitrified bulk solvent suggests there is substantial contraction of the solvent on cooling. Both the side chain ordering and solvent contraction may drive the unit cell contraction. In either case, the observed temperature dependence of intermolecular interactions suggests that caution may be necessary in interpreting protein-protein and protein-nucleic acid interactions based on low-temperature crystal structures.

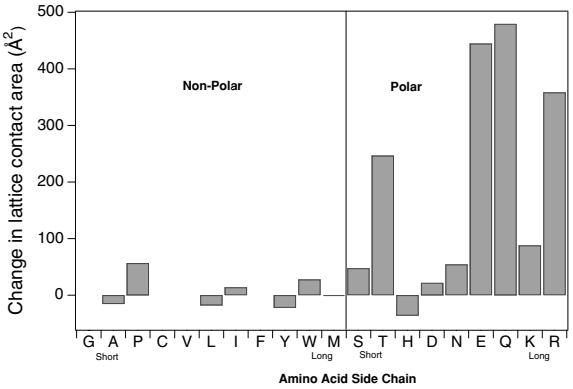

Figure 1. Change in lattice contact area with cryocooling. Overall, the lattice contacts increase in area by %50. This increase is dominated by the long, polar side chains. Although other types of side chains are on the surface (data not shown), they are not as involved in the contact change.

REFERENCES

1. R.H. Jacobson, X.J. Zhang, R.F. DuBose, and B.W. Matthews, Nature **369**, 761 (1994).
2. D.H. Juers, R.H. Jacobson, D. Wigley, X.J. Zhang, R.E. Huber, D.E. Tronrud, and B.W. Matthews, Protein Science **9**: 1685 (1999).

Structural Studies of a Full-Length DEAD-box Protein

Andrew B. Carmel and Brian W. Matthews
Institute of Molecular Biology, Departments of Chemistry and Physics,
and Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403

DEAD-box proteins are a ubiquitous class of proteins found in organisms from bacteria to humans. DEAD-box proteins are thought to be RNA helicases that use the energy from ATP hydrolysis to unwind short stretches of duplex RNA in processes requiring RNA structure rearrangement. These proteins contain a core domain consisting of eight highly conserved sequence motifs, flanked by extensions of various lengths on the N- and C-termini. While DEAD-box proteins possess RNA-dependent ATPase activity and *in vitro* studies have identified some members having ATP-dependent RNA helicase activity, the *in vivo* roles of these proteins have yet to be determined.

We have crystallized a full length protein which contains all of the eight conserved motifs found in DEAD-box proteins. We collected a single wavelength data set on Beamline 5.0.2 with the resolution limit extending to 3.7Å. This was a vast improvement in data resolution and quality over data collected on our home (RAXIS IV) source. A much lower resolution (~6.0Å) data set was collected at multiple wavelengths using crystals of selenomethionine substituted protein. However, due to the low resolution limit, this data was not useful for MAD phasing. The protein crystallizes in space group $P4_12_12$ or $P4_32_12$ with unit cell dimensions a=b=100.3Å, c=110Å.

Structure Determination of P22 lysozyme: A Possible Evolutionary Link between Two Families of Lysozymes

B. H. M. Mooers, N. C. Gassner, M. M. Dixon,
R. F. DuBose, X.-J. Zhang, W. A. Baase, and B. W. Matthews.
Institute of Molecular Biology, Department of Physics,
and Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403

The product of gene *19* from the *Salmonella* phage P22 is essential for the lysis of the bacterial cell wall. Its substrate specificity is very similar to that of phage T4 lysozyme with which it has 26% sequence identity. However, in common with goose egg-white lysozyme, P22 lysozyme lacks two segments of amino acids which are surface loops in T4 lysozyme. These similarities and differences suggest that P22 lysozyme may provide an evolutionary link between T4 lysozyme and goose egg-white lysozyme[1]. To determine the similarity of the structure of P22 lysozyme relative to other lysozymes, we have crystallized P22 lysozyme and collected 3.0 Å data in house. Multiple approaches to the structure determination of P22 lysozyme by molecular replacement have been unsuccessful. We now have crystals of the protein with selenium-methionine incorporated for a MAD experiment. We were able to collect 2.9 Å diffraction data of poor quality with cryo conditions at BL 5.2.1. We are working on improvements of the crystals and the cryo conditions in order to enhance the quality and resolution limit of the diffraction data.

REFERENCE

1. L. H. Weaver, D. Rennell, A. R. Poteete, and Matthews, B. W. *J. Mol Biol*. **184**, 739-741 "Structure of phage P22 gene 19 lysozyme inferred from it homology with phage T4 lysozyme." (1985).

Automated Redesign of T4 Lysozyme

B. H. M. Mooers[1,3,4], D. Datta[5], W. A. Baase[1,3,4], S. L. Mayo[5], and B. W. Matthews[1,2,3,4].
Institute of Molecular Biology[1], Department of Physics[2], and
Howard Hughes Medical Institute[3], University of Oregon, Eugene[4], OR 97403 and
California Institute of Technology[5], Pasadena, CA 91125

Recent successes in the automated design of small proteins suggest that these methods may be useful in addressing the protein stability problem. The net difference between the free energy of the folded and unfolded forms of proteins is small and can be substantially altered by changing a single amino acid. If the energy contributions associated with an amino acid change can be accurately modeled by automated design methods, then the structures predicted by these methods can provide a rational basis for experiments that are designed to further test our understanding of protein stability. In collaboration with the Steve Mayo and his group at Cal Tech, we are extending automated design methods to studies of the thermal stability of bacteriophage T4 lysozyme, a very well characterized mid-size globular protein with over 1100 mutant proteins and 300 mutant crystal structures known. Starting with the crystal structure of the wild-type (WT) enzyme, we used the automated design method developed by Mayo's group to find the amino acid sequences and side-chain conformations of twenty-four residues in the core of the C-terminal domain of T4 lysozyme which pack well with the surrounding backbone and side-chain atoms. Two sequences were found (one with seven residue changes and one with ten residue changes) which were predicted to be more stable than the wild-type protein. These two multiple mutant proteins, selected subsets of these multiple mutants, and the single mutants, in the form of sequences, three-dimensional structures, and relative stabilities, are being tested against known T4 lysozyme mutants or by construction of selected variants. The rigorous and objective testing of automated protein design algorithms with T4 lysozyme will suggest the strengths, limitations, possible improvements, and degree to which such procedures may be applied to globular proteins, including those of biomedical importance.

Accurate crystal structures of the multiple and single mutants are required to evaluate the these predictions. High resolution diffraction data for most of these mutants are being collected in-house with cryocrystallographic techniques; however, some difficulty has arisen in obtaining sufficiently large crystals of every mutant. One case was the double mutant V149I/T152V. We were able to collect 1.52 Å data of this mutant at BL-5.2.1. Crystals of the mutant protein are in space group $P3_221$ and have unit cell dimensions that are isomorphous with that of the wild type protein. The unit cell lengths were a=b=59.6 Å and c=95.3 Å. The data set is 93.5% complete and has a Rsym of 5.8%. The current model has a R-factor of 17.4% and will be useful in the evaluation of the predicted structures.

Funding Sources: NIH Grant GM20066 and HHMI.

Principal Investigator: Brian Matthews, Institute of Molecular Biology, Department of Physics, and Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403. Email: brian@uoxray.uoregon.edu. Telephone: (541) 346-2572.

Contact Person: Andrew Carmel, Institute of Molecular Biology, Departments of Chemistry and Physics, and Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403. Email: carmel@uoxray.uoregon.edu. Telephone: (541) 346-4080.